

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

TITLE OF THE INVENTION

METHOD AND APPARATUS FOR DISCOVERING NETWORK TOPOLOGY

INVENTORS

Priya Govindarajan
a citizen of India,
residing at 3120 NW John Olsen Ave., #25-301,
Hillsboro, OR 97124

David M. Durham
a citizen of the United States,
residing at 1024 NE Parksedge Cir.
Hillsboro, OR 97124

Prepared by

BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN LLP
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1026
(303) 740-1980

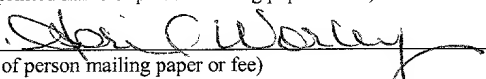
EXPRESS MAIL CERTIFICATE OF MAILING

"Express Mail" mailing label number: EL845313337US

Date of Deposit: March 30, 2001

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Commissioner of Patents and Trademarks, Washington, D. C. 20231

April M. Worley
(Typed or printed name of person mailing paper or fee)


(Signature of person mailing paper or fee)

3/30/01
(Date signed)

METHOD AND APPARATUS FOR DISCOVERING NETWORK TOPOLOGY

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention relates generally to the field of computer network management. More particularly, the invention relates to determining the topology of a network using devices distributed in different locations on the network.

Description of the Related Art

[0002] Any networked domain will typically have a number of devices responsible for routing the traffic around the network. Typically, the devices are routers or computers that support layer 2 (data link layer of the ISO's (International Standards Organisation's) OSI (Open System Interconnection) standard for networking) and layer 3 (network layer of OSI). In order to survive a failure of any one device or link, a network is typically designed with redundant back-up routes. Some networks can have multiple interconnected back-up routes. Over time, devices can be moved, added, removed or replaced, routes can be moved or altered and device configurations can be changed. As a result, any documentation of the original network design, if it was created, becomes outdated. When problems occur, troubleshooting is made more difficult without accurate information on the network's topology. When new demands are to be applied to the network, it may not be possible to determine how to reconfigure the network without knowledge of the network's existing topology. When performance is to be measured or improved, it is difficult to determine how

to develop the network without understanding all of the links and which devices are connected to each one.

[0003] Some devices are supplied with mechanisms that allow them to determine their connections. These mechanisms might allow someone to determine the connections at each device, compile the information together and, from there, determine the network's design. However, these mechanisms typically require that all of the devices use the same mechanism and do not enable the devices to communicate with each other to compile the information.

[0004] PING (Packet Internet Groper) is supported by a large number of different network devices. For those systems that do not support PING, there are many freeware and shareware utilities available for many different platforms that will support PING. PING sends an ICMP (Internet Control Message Protocol) packet to a specified address and waits for a reply. If a reply is received, then a connection to the addressed device is confirmed. However, PING does not determine connections to devices for which it does not know an address nor does PING determine the routing that was used to get to a device.

[0005] Traceroute is a utility that is available in many forms to operate on many different devices. UNIX and most Windows operating systems include a Traceroute utility. It is typically used to trace out packet routes through Internet connections. Traceroute traces the route that a packet sent from one device to another takes, shows how many hops the packet takes, and identifies the device at each hop. The identified devices at each hop are typically the routers. Traceroute operates by sending packets with low TTL (time-to-live) fields. When the packet does not reach its destination because the number of hops exceeds the TTL value, the last host that the packet did reach returns the packet and identifies itself. Traceroute sends a series of packets and increments the TTL of each one. Each failed packet is returned with an identification of the next host. By sending the

[illegible]

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0006] The appended claims set forth the features of the invention with particularity. The present invention is illustrated by way of example, and not by way of limitation, in the figures of the accompanying drawings in which like reference numerals refer to similar elements and in which:

[0007] Figure 1 is a block diagram of a computer network suitable for implementing the present invention;

[0008] Figure 2 is a flow chart showing a process for discovering the topology of a network according to one embodiment of the invention; and

[0009] Figure 3 is an example of a typical computer system of the network of Figure 1.

DETAILED DESCRIPTION OF THE INVENTION

[0006] The present invention can be implemented by using PING and Traceroute at sources distributed throughout a network and then piecing all of the collected information together to determine the overall topology of the network. The PING and Traceroute sources will be called Smartlinks. They respond to policies sent to them from a designated central server. In brief, the Smartlinks register themselves to the server and the server then sends policies to each of the Smartlinks. The Smartlinks execute the policies by running Traceroutes to each other. This allows them to identify all of the hosts that are between each Smartlink. This information is then sent to the central policy server which compiles the reports from each Smartlink to determine the topology of all of the network that is within the Smartlinks. The server can then render the topology in the form of lists, charts, graphs or a database.

[0007] Figure 1 shows an example of a network that has redundant paths. The network is shown as populated with routers and Smartlinks and delimited by end hosts. It may also include a number of other devices not shown. The network of Figure 1 is a simplified example of the type of network that can benefit from implementing the present invention. The present invention can be applied to networks of myriad different configurations and topologies. In one embodiment, the network devices are common commercially available Intel Pentium® II and III - based microcomputers running a Windows NT™ operating system. The microcomputers are connected in this embodiment by Ethernet and use IP (Internet Protocol) messaging. The network can also support SNMP (Simple Network Management Protocol) and RMON (Remote Monitoring) network management. However, the nature of the network devices, the nature of the network connections and the nature of the network protocols may be different in order to suit the demands of particular

applications. The messaging and utilities discussed in the examples provided below are supported by many different hardware and software platforms and can easily be adapted for many more.

[0008] In Figure 1, the network is delimited by three end hosts 10, 12, and 14. Between the end hosts are four routers 18, 20, 22, 24. Each router has at least two ports that are independently addressable and oriented toward different directions in the network. The first router is coupled to the first end host 10 at one port. It is coupled to the second and third routers respectively, at its other two ports. The second router 20 is connected to the first router 18 at one port and to the fourth router 24 at its other port. The third router 22 is connected to the first router 18, the second end host 12 and the fourth router 24 at its respective three ports. Finally, the fourth router 24 is connected to the second router 20, the third router 22 and the third end host 14 at its respective three ports. As can be seen in Figure 1, messages from, for example, the first end host can be sent to the third end host through either the second or third router so that there is redundancy in the event that one of the paths should fail. The end hosts may be connected to other devices, to the Internet, an intranet or to some other WAN (wide area network). As mentioned above, the network configuration shown in Figure 1 is provided as a working example only and the invention is not limited to such a network topology, scale or size.

[0009] The network of Figure 1 has been augmented by the insertion of four Smartlinks 26, 28, 30, 32. The specific placement and number of Smartlinks can be selected based on many different criteria. More thorough results may be obtained if the Smartlinks are at the edges of the network, if the edges are known. Less traffic may be required to determine the network topology when the Smartlinks are not adjacent to each other. However, the network can also be discovered without any knowledge of the network topology and by placing the Smartlinks in random locations. The number of Smartlinks can be selected to strike a balance between cost and speed. Quicker

results can be obtained with at least two Smartlinks. An overly large number of Smartlinks may burden the network with too many redundant messages. An overly small number of Smartlinks may require too many messages to determine the complete topology of the network. The Smartlinks may be unique devices dedicated to implementing the Smartlink functionality or network devices already present on the network for other purposes including routers. A unique Smartlink device can be constructed from a conventional programmable network appliance such as the Intel Pentium® II or III - based microcomputers mentioned above.

[0010] In the example of Figure 1, the Smartlinks each have two ports and are inserted into the network connections described above. The first Smartlink 26 is connected between the first end host 10 and the first router 18. The second Smartlink 28 is connected between the second end host 12 and the third router 22. The third Smartlink 30 is connected between the second router 20 and the fourth router 24. The fourth Smartlink 32 is connected between the third end host 14 and the fourth router 24. These locations are chosen to place the Smartlinks at the edge of the network, near the end hosts and spaced apart from each other. The Smartlinks initiate their network discovery process and report their results to a server which then compiles the results to determine the network topology. The server may be one of the three end hosts or one of the Smartlinks or another network appliance not shown in Figure 1. For purposes of the present example, the server will be selected to be the first end host 10.

[0011] The Smartlinks can be inserted into the network without disturbing the normal operation of the network. The steps described below rely on IMCP (Internet Control Message Protocol) but can be adapted to any other protocol. For an ICMP-compliant network, no knowledge of any device, its configuration or how it operates is required. In particular, no MIB (Management Information Base) is required. Compatibility issues are completely avoided. The Smartlinks can be

expanded functionally to perform other tasks, such as analyzing the capability and capacity of the network links to which they are connected and for traffic engineering.

[0012] Referring to Figure 2, the discovery of the network's topology begins, in one embodiment, with the Smartlinks 26, 28, 30, 32 which start up 42 and then register themselves 44 with the server 10. The server which has already been started up 40 receives the registrations 44 from each of the Smartlinks. In their registrations, the Smartlinks identify their IP (Internet Protocol) addresses, their port addresses (physical addresses) and an identification code. In the present example, each Smartlink has two network interfaces. This allows a Smartlink to be inserted into the network at any location. However, the number of network interfaces or ports is not important to the invention, more or fewer can be provided.

[0013] Using the registration information, the server sends Traceroute policies to each of the Smartlinks. Having the Smartlinks register themselves on startup, allows Smartlinks to be inserted into the network or taken out at any time without entering any information about the reconfiguration into the server. It also ensures that only Smartlinks capable of communicating with the server will be registered. A Smartlink coupled to a failed connection will not be able to register and the resulting topology will show a gap or end where that Smartlink is located. It is possible for all network devices to be programmed to register with the server. This can aid the server in discovering the network but will not reveal where the network connections are. In the present example, only the Smartlinks register with the server.

[0014] The Traceroute policies are addressed to each Smartlink using conventional protocols from the policy server. The policies instruct each Smartlink to run Traceroute messages to specified destinations. In one embodiment, each Smartlink is instructed to run Traceroute to each of the other Smartlinks. Typically the policies will include IP addresses for each of the Smartlinks. Further

policies can also be sent, including the timing and order in which the messages are to be sent, the identity and addresses of other devices on the network, and the protocols or utilities to be used in sending the messages. While the present invention is described in terms of running PING and Traceroute utilities that are offered with Windows NT, other utilities can be used. Alternatively, no utility need be used, the Smartlinks can strategically send messages and wait for replies based on the network's protocols. In any event, the information will be gathered by polling all the devices between the Smartlinks and discovering how they are all connected together.

[0015] After the Smartlinks receive the policies 46, they run PING and Traceroute messages 48. PING is used to save time in determining how to direct the Traceroute packets. Consider, for example, Smartlink 1. The policy will direct it to run a Traceroute message to Smartlinks 2, 3 and 4. Starting with Smartlink 2 it must send its Traceroute message either up through one network interface or port or to the right through the other network interface as shown in Figure 1. In one embodiment, Smartlink 1 will do both and wait for both results. However, the Traceroute message is slower and takes more system resources than a PING. So, in another embodiment, Smartlink 1 sends a PING message out both ports in both directions. The upward PING goes to the first end host 10 where it cannot be sent to any other devices and is not answered. The PING to the right, goes through the routers 18, 22 and any other devices on the network until it reaches Smartlink 2, where an answer is generated and returned to Smartlink 1. Smartlink 1 has now determined the network interface through which it should send its Traceroute message to begin discovering the connections between all the devices.

[0016] Smartlink 1, will also PING for Smartlinks 3 and 4. Smartlink 3 can be reached by Smartlink 1, both through the second router 20 and the fourth router 24, depending on the path that the PING message takes. In some networks, it may be possible for the PING to travel both paths and

Smartlink 1 will receive two replies one from each port. Smartlink 3's reply will identify Smartlink 3 and the network interface of Smartlink 1 that sent the PING. This gives Smartlink 1, not only the direction to send messages to reach Smartlink 3 but also the knowledge that Smartlink 3 is accessible to it.

[0017] After PINGing the other Smartlinks, Smartlink 1 will send Traceroute messages to each of the other Smartlinks or just to the Smartlinks that the server named in its policies. As mentioned above, Traceroute is selected as a convenient way of implementing the functions that are desired to show the data paths, however, other utilities that offer the important desired functionality can be used instead. The Traceroute utility will send a message to the destination Smartlink with a low TTL (time to live) and then increment the TTL until the path is traced through the target. For the example of Smartlink 1 tracing the route to Smartlink 2, a message with a TTL of 1 will first leave the port facing the first router and expire at the first router. The first router will respond to the message, send an error report and identify itself. It will also identify an IP address of the router at the port that received the message. The message that is returned is typically a standard error message that, according to ICMP, contains all of the information that the Traceroute utility desired to obtain.

[0018] The second Traceroute message will be sent to Smartlink 2 with a TTL of two. It will make it to the third router. The message will expire at the third router after making the two hops that it takes to get there. The third router ICMP will send an error message back to Smartlink 1 identifying itself and an IP address of the port that received the message.

[0019] The third Traceroute message with TTL of three makes it to Smartlink 2 in three hops. Smartlink 2 will accordingly reply and identify itself and the IP address of the port at which the message was received. Smartlink 1 has now identified the three hops that it takes to get to Smartlink 2. It collects the Traceroute results and puts it into a table or database that it can send to

the server. In Windows NT, the Traceroute utility, called Tracert, provides the IP address of each device on the way to Smartlink 2 and the order in which that address was received in an error message. Instead of a database or table, Smartlink 1 can simply send this list. The result of the Traceroute runs will depend upon the utility that is selected. The format of the report of the server can be selected to optimize efficiency for the system involved. The Traceroute utility is then repeated to discover the routes to the other Smartlinks.

[0020] At the same time that Smartlink 1 is tracing the path to Smartlink 2, Smartlink 2 is tracing the path to Smartlink 1. It will find that its network interface facing the third router gets an answer from Smartlink 1 to its PING and that there are three hops from it to Smartlink 1. The hops are the third router, the first router and Smartlink 1 itself. While the results of the two Traceroute runs involve the same devices, the results are not necessarily the same because different ports are used. The list of IP addresses collected by the two Smartlinks can, accordingly, be different. Smartlinks 1 and 2 will each compile a table or database of results for the same equipment but the IP addresses may be very different.

[0021] After each Smartlink has completed all of its PING and Traceroute runs, it sends its results as a table, database, list or any other form to the server 48. The server collects the results from each Smartlink and makes a router list 50. To complete the topology, the server should correlate the different IP addresses to each router. In one embodiment, the routers support SNMP (Simple Network Management Protocol) or RMON (Remote Monitoring) and maintain a MIB (Management Information Base) that lists all of the IP addresses of the router. Accordingly, the server can query each identified router to provide a list of its IP addresses from its MIB. If the router has no MIB or the format of the MIB is not known to the server, then the server can correlate the IP addresses by sending packets with non-existent ports 52.

[0022] The server, by sending a packet to a router, using an IP address of the router but addressed to an application port that does not exist on the router will provoke an error message from the router. According to ICMP, the error message will identify, the router, the address of the interface of the router that received the message and is responding to the message, and the message itself. The server sends these packets to all of the routers for all of the IP addresses and collects the data in the error messages 54. The error messages allow the server to correlate different IP addresses of the same router to that same router. In doing so, the server collects a list of all of the source IP addresses for each router which it can then correlate to the Traceroute results.

[0023] As described above, the present invention discovers the connections of the routers. It can also rely on MIBs or similar types of tables and directories to discover any other devices. The end hosts are not discovered. The routers are discovered easily because they respond to Traceroute utilities and provide rich information under ICMP. However, the process can be adapted to provide information about other devices on the network with suitable modifications to the messages and the replies. The information that is discovered by sending messages and replies can be supplemented with any other reliable source, such as MIBs. The resulting data provides a list of all of the IP addresses of each router and all of the connections between ports as identified by IP address. From this data, a drawing, table, chart, spreadsheet or any other rendering of the network topology can be generated 60. Figure 1 is an example of such a chart.

[0024] A computer system 400 representing an example of a system upon which features of the present invention may be implemented is shown in Figure 3. The network appliances of Figure 1, including the Smartlinks, will typically be configured similarly to what is shown in Figure 3. The computer system 400 includes a bus or other communication means 401 for communicating information, and a processing means such as a microprocessor 402 coupled with the bus 401 for

processing information. The computer system 400 further includes a main memory 404, such as a random access memory (RAM) or other dynamic data storage device, coupled to the bus 401 for storing information and instructions to be executed by the processor 402. The main memory also may be used for storing temporary variables or other intermediate information during execution of instructions by the processor.

[0025] The computer system may also include a nonvolatile memory 406, such as a read only memory (ROM) or other static data storage device coupled to the bus for storing static information and instructions for the processor. A mass memory 407 such as a magnetic disk or optical disc and its corresponding drive may also be coupled to the bus of the computer system for storing information and instructions.

[0026] The computer system can also be coupled via the bus to a display device or monitor 421, such as a cathode ray tube (CRT) or Liquid Crystal Display (LCD), for displaying information to a user. For example, graphical and textual indications of installation status, operations status and other information may be presented to the user on the display device. Typically, an alphanumeric input device 422, such as a keyboard with alphanumeric, function and other keys, may be coupled to the bus for communicating information and command selections to the processor. A cursor control input device 423, such as a mouse, a trackball, or cursor direction keys can be coupled to the bus for communicating direction information and command selections to the processor and to control cursor movement on the display 421.

[0027] A communication device 425 is also coupled to the bus 401. The communication device 425 may include a modem, a network interface card, or other well known interface devices, such as those used for coupling to Ethernet, token ring, or other types of physical attachment for purposes of providing a communication link to support a local or wide area network (LAN or WAN),

such as those used for coupling to Ethernet, token ring, or other types of physical attachment for purposes of providing a communication link to support a local or wide area network (LAN or WAN), for example. In this manner, the computer system may also be coupled to a number of clients or servers via a conventional network infrastructure, including an intranet or the Internet, for example.

[0028] It is to be appreciated that a lesser or more equipped computer system than the example described above may be preferred for certain implementations. Therefore, the configuration of the exemplary computer system 400 will vary from implementation to implementation depending upon numerous factors, such as price constraints, performance requirements, technological improvements, or other circumstances.

[0029] It should be noted that, while the steps described herein may be performed under the control of a programmed processor, such as the processor 402, in alternative embodiments, the steps may be fully or partially implemented by any programmable or hard coded logic, such as Field Programmable Gate Arrays (FPGAs), TTL logic, or Application Specific Integrated Circuits (ASICs), for example. Additionally, the method of the present invention may be performed by any combination of programmed general purpose computer components or custom hardware components. Therefore, nothing disclosed herein should be construed as limiting the present invention to a particular embodiment wherein the recited steps are performed by a specific combination of hardware components.

[0030] In the present description, for the purposes of explanation, numerous specific details are set forth in order to provide a thorough understanding of the present invention. It will be apparent, however, to one skilled in the art that the present invention may be practiced without some of these specific details. In other instances, well-known structures and devices are shown in block

diagram form. The specific detail may be supplied by one of average skill in the art as appropriate for any particular implementation.

[0031] The present invention includes various steps, which may be performed by hardware components or may be embodied in machine-executable instructions, such as software or firmware instructions. The machine-executable instructions may be used to cause a general-purpose or special-purpose processor programmed with the instructions to perform the steps. Alternatively, the steps may be performed by a combination of hardware and software.

[0032] The present invention may be provided as a computer program product that may include a machine-readable medium having stored instructions thereon, which may be used to program a computer (or other machine) to perform a process according to the present invention. The machine-readable medium may include, but is not limited to, floppy diskettes, optical disks, CD-ROMs, and magneto-optical disks, ROMs, RAMs, EPROMs, EEPROMs, magnet or optical cards, flash memory, or any other type of medium suitable for storing electronic instructions. Moreover, the present invention may also be downloaded as a computer program product, wherein the program may be transferred from a remote computer to a requesting computer by way of data signals embodied in a carrier wave or other machine-readable propagation medium via a communication link (e.g., a modem or network connection).

[0033] Importantly, while embodiments of the present invention are described with reference to discovering the topology of servers coupled on a network communicating using ICMP, the method and apparatus described herein are equally applicable to any other type of network and to other communications protocols. In addition, while the invention has been described in terms of a wired computer network it can also be applied to wireless networks that employ point-to-point and point-to-multipoint connections.

[0034] Although this disclosure describes illustrative embodiments of the invention in detail, it is to be understood that the invention is not limited to the precise embodiments described. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense. Various adaptations, modifications and alterations may be practiced within the scope of the invention defined by the appended claims.